(1)

ARI Research Note 90-112

AD-A227 059

The Object Display in Decision Making and Judgment

Barbara Barnett, Barry P. Goettl, Arthur F. Kramer, and Christopher D. Wickens

University of Illinois

for

Contracting Officer's Representative
Michael Drillings

Basic Research
Michael Kaplan, Director

August 1990

FIFTY YEARS OF SERVICE
1940 - 1990

United States Army
Research Institute for the Behavioral and Social Sciences

90 09 21 052

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Research accomplished under contract for the Department of the Army

University of Illinois

Technical review by

Nehama Babin

NOTICES

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS —
2a. SECURITY CLASSIFICATION AUTHORITY — 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE —	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CPL 86-7	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARI Research Note 90–112
6a. NAME OF PERFORMING ORGANIZATION Department of Psychology University of Illinois 6b. OFFICE SYMBOL (If applicable) —	7a. NAME OF MONITORING ORGANIZATION U.S. Army Research Institute for the Behavioral and Social Sciences
6c. ADDRESS (City, State, and ZIP Code) 603 East Daniel Street Champaign, IL 61820	7b. ADDRESS (City, State, and ZIP Code) 5001 Eisenhower Avenue Alexandria, VA 22333-5600
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Research Institute 8b. OFFICE SYMBOL (If applicable) PERI-BR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA903-83-K-0255
8c. ADDRESS (City, State, and ZIP Code) 5001 Eisenhower Avenue Alexandria, VA 22333–5600	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: 61102B PROJECT NO.: 74F TASK NO.: NA WORK UNIT ACCESSION NO.: NA

11. TITLE (Include Security Classification)
The Object Display in Decision Making and Judgment

12. PERSONAL AUTHOR(S)
Barnett, Barbara; Goettl, Barry P.; Kramer, Arthur F.; and Wickens, Chistopher D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 83/08 TO 86/09	1990, August	22

16. SUPPLEMENTARY NOTATION
Michael Drillings, contracting officer's representative.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Decision making, Perception, Attention, Graphs, Information, Integration, Object integration, Judgment

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper presents three experiments which apply the principle of compatibility of proximity to multicue decision making and judgment. This principle states that whenn a task requires the integration of information, the display should present the data in close proximity. When, however, the information must be selectively attended or processed, the information should be displayed with low proximity. In the first experiment, subjects processed information cues in a multicue decision making task. The cues varied in their source, diagnosticity and reliability. They were presented either as separate bargraphs, or as rectangles. Their presentation was either simultaneous or sequential. The results indicated that greater proximity in time (simultaneous) and objectness (rectangular cues) lead to improved judgments. Furthermore, the rectangular format did not disrupt the memory for isolated cue values. In experiments 2 and 3, subjects were shown three numerical values and were required to combine them to predict a criterion value. Both the equation relating the cues to the criterion and the display format of the cues were varied. Experiment 2 demonstrated that when the cues had to be

Continued

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Michael Drillings	22b. TELEPHONE (Include Area Code) (202) 274–8722 22c. OFFICE SYMBOL PERI-BR

DD FORM 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

ARI Research Note 90-112

19. Abstract (Continued)

selectively attended, performance was better with a bargraph display than with an integrated triangle display. No difference was observed when all the cues had to be integrated. Experiment 3 demonstrated that the effects of display format are contingent upon the relation between the cues and the criterion. Discussion centers on display format and perception of relations between numerical data.

The Object Display in Decision Making and Judgment

Barbara Barnett, Barry P. Goettl, Arthur F. Kramer
and Christopher D. Wickens

Institute of Aviation
University of Illinois at Urbana-Champaign
Savoy, Illinois

INTRODUCTION

Most decision making requires information integration. Various cues, often of differing degrees of reliability, must be processed and compared to form an aggregate diagnosis, or estimate of the present or future state of the world, upon which an action should be based. Large amounts of experimental data exist on the optimal cognitive strategies for integrating such information, and on human departures from or adherence to these strategies (e.g., Anderson, 1981; Lopes, 1981; Pitz & Sachs, 1984). There is an equally ample data base bearing on the appropriate sensory and perceptual factors to be considered in display design (e.g., Boff, Kaufman, & Thomas, 1986). However, little research exists regarding the optimal display interface between multiple information channels and the appropriate mental integration process. We present here three experiments that test a principle of compatibility of proximity, which asserts that, to the extent that information must mentally be integrated, such information should be physically integrated or proximate as well (Carswell & Wickens, in press; Casey & Wickens, 1986; Wickens, 1986). Proximity may be defined by closeness in space, but also in terms of other shared stimulus attributes, objectness, or other Gestalt principles of grouping. The first experiment examines information integration in the context of a dynamic, sequential multicue decision making task. The second and third experiments are each variations on a classic multicue probability learning task.

EXPERIMENT 1: MULTICUE JUDGMENT

Experiment 1 investigates the compatibility of proximity advantage by manipulating orthogonally three proximity dimensions: space, time, and the degree to which two cues are integrated into a single object (i.e., two bargraphs vs. height and width of a rectangle).

Support for the advantage of object integrality in information integration has previously been demonstrated by a number of studies in our laboratory. Carswell and Wickens (in press) found an integrality advantage using a triangle display, while Jones and Wickens (1986) demonstrated the same advantage using a pentagon display. While both of these investigations were carried out in a process control environment, another experiment using a multicue decision making task, demonstrated improved performance with a rectangle display relative to a digital display (Wickens & Scott, 1983).

In the context of medical diagnosis, Cole (1986) represented the total volume of air exchanged by the ventilator or the patient in a fixed period of time as the area of a rectangle. For each rectangle, the width represented mean breathing rate for that period, while mean volume per breath was represented by its height.

The particular value of the rectangle in the context of the display used by Cole (1986), by Wickens and Scott (1983), and in the present experiment, is that the two dimensions represented combine multiplicatively to define a quantity of direct importance to the operator. In the present study, this quantity is the total information worth of the cue--defined by the product of reliability and diagnosticity (Johnson et al., 1973). To the extent that area is a commodity that can be perceived directly, then it

serves as an "emergent feature" which helps to facilitate the integration task at hand (Pomerantz, 1981).

While the benefits of the integral display formats have been demonstrated, these have not been achieved without cost in many instances. This cost has been manifested in the sacrifice of processing individual cue attributes, while focusing on the more global, integral "whole" of the system. Such a performance decrement in processing isolated attributes or applying more focal attention on the integral displays has also been demonstrated by a number of studies in our laboratory (Carswell & Wickens, 1986; Casey & Wickens, 1986; Goettl, Kramer & Wickens, 1986; see Wickens, 1986 for an integrative summary of these data).

Based on findings of previous research and on the notions of compatibility of proximity, it was our hypothesis that the information integration task would be facilitated by the more integral display formats. Furthermore, it was hypothesized that memory for isolated cue attributes would be harmed by such display integrality. Finally, it was hypothesized that conditions showing closer proximity in space and/or time will increase the ability to perform this object integration task for all three levels of object integrality.

Method: Experiment 1

Subjects. Twenty-one males and 19 females participated in this study. The participants were students from the University of Illinois and were paid for their efforts. Bonuses were offered as performance incentives.

Experimental Scenario. Subjects were asked to envision themselves as military aircraft pilots, deciding in mid-flight whether to continue or abort the current mission. This decision was based on weighted evidence from a number of cues, presented by an intelligent on-board decision aid. The cues provided diagnostic information as to the state of the aircraft, as well as information about weather conditions, enemy strength and navigational equipment status. Every trial consisted of the presentation of a subset of four out of eight possible cues, each varying in total information worth. All eight possible cues had a unique information worth that remained constant across trials and blocks. Information worth, or the probability that the cue's value depicts the actual state of the world, was defined as the multiplicative combination of a cue's reliability (The accuracy with which a displayed cue conveys the true state of the cue) and diagnosticity (How informative each cue is in evaluating the potential success or failure of the mission).

Cue values were dichotomous in that the information presented supported one of only two possible hypotheses: continuing the mission or aborting the mission. These values, weighted by each cue's information worth, were to be combined according to the following algorithm:

$$\text{Opt dec} = (\sum W\ C - \sum W\ A) \times .137 \quad (1)$$

where W C is the sum of the weighted cues in support of continuing the mission, W A is the sum of the weighted cues in support of aborting the

mission, and .137 is a constant, converting scores to the appropriate response scale values.

Design. Display integrality was manipulated between three groups of eight subjects. In the bargraph group, subjects saw reliability and diagnosticity values for each cue represented as the heights of two bars side-by-side on a single graph (Figure 1, left). Subjects in the rectangle group saw reliability and diagnosticity values for each cue represented respectively as the width and height of a single rectangle (Figure 1, center). Subjects in the integral rectangle condition also saw four rectangles representing reliability and diagnosticity values of the four cues (as in the previous condition), however the four rectangles were joined at the innermost corner to form one integral geometric figure (Figure 1, right). The dichotomous cue values were represented as shaded (continue the mission) or unshaded (abort the mission) figures.

Proximity in space and time was varied at four levels within subjects. Each of the four presentation conditions are described as follows:

1) Separate in Space and Time: Four cues presented sequentially (one second per cue) at separate locations on the screen. This condition minimizes the degree of proximity.

2) Separate in Time: Four cues presented sequentially (one second per cue) in the center of the screen (i.e., at the same location).

3) Simultaneous, Speed Stress: Four cues presented simultaneously for one second in separate locations on the screen.

4) Simultaneous, No-Speed Stress: Four cues presented simultaneously for four seconds in separate locations on the screen. This condition maximizes the degree of integrality of proximity.

The manipulation of speed stress had two functions. One was to determine the susceptibility of the three kinds of displays to stress. The second was to provide two anchors for the control of timing, appropriate to compare with the separate conditions. The speed stress condition presented the simultaneous cues for as long as each cue was presented in the sequential condition (1 second). The no-speed condition presented the 4 cues for as long as all cues were presented in the sequential condition (4 seconds).

In all of the above conditions, subjects were called upon to make two kinds of responses. The primary task response was to indicate the aggregated weight of evidence for one hypothesis or the other. As a secondary response, subjects were probed at random intervals as to the value of particular variables. The purpose of this secondary task was to establish the extent to which memory for isolated values may have been influenced by display proximity.

Following presentation of each set of four cues, the subject had 10 seconds in which to respond by positioning a marker, controlled by a joystick, on the linear scale shown at the bottom of Figure 1. Depressing a

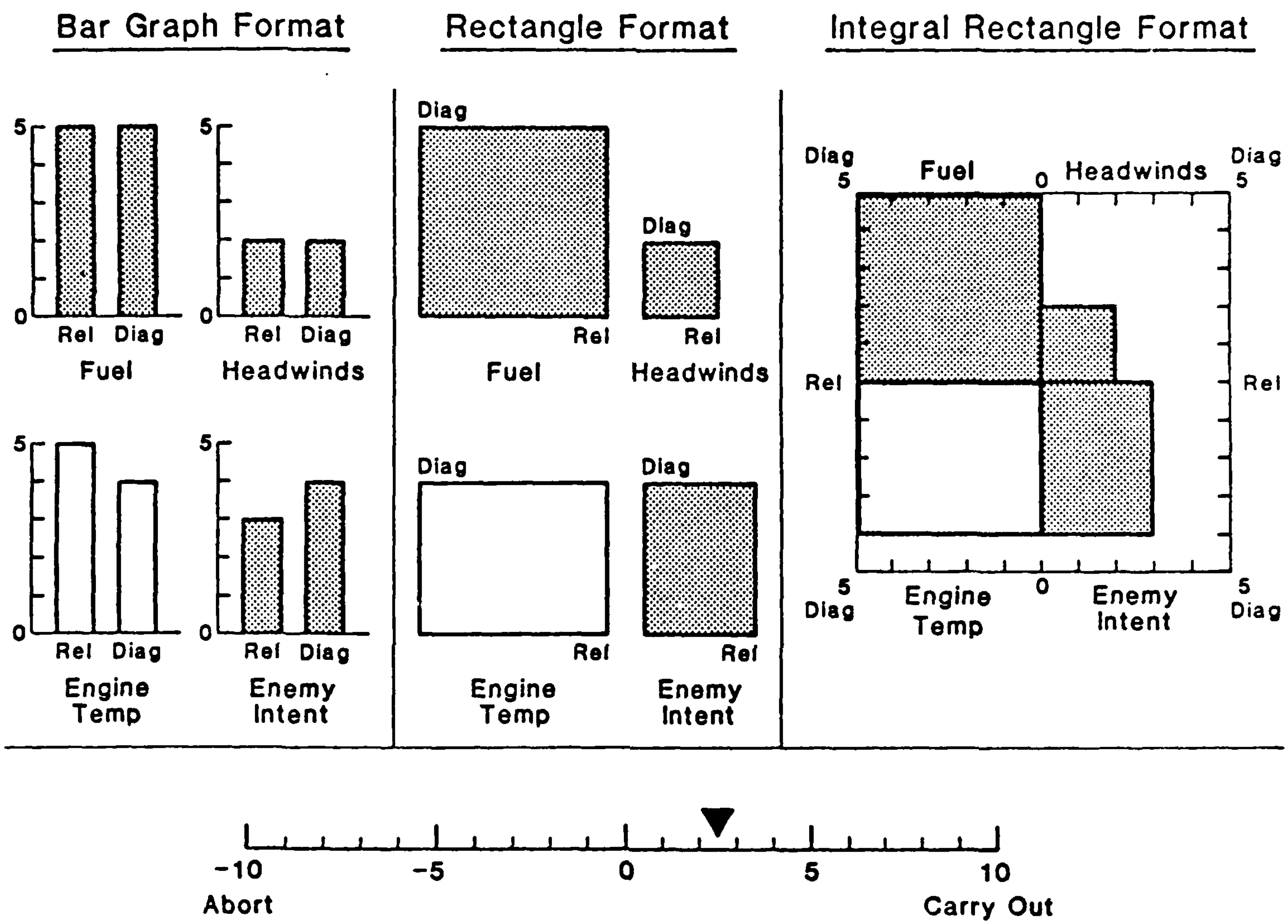

Figure 1. The three display integrality formats, simultaneous conditions. Only four of the eight possible cues are displayed. The same four positions were used for the separate in time (sequential) and space condition. The analog response scale used for all three formats is displayed below (experiment 1).

button on the joystick then entered the subject's response. Subjects participated in one practice session (consisting of four practice blocks of 40 trials--one block for each of the proximity conditions), followed by two experimental sessions. Each experimental session consisted of two blocks of 80 trials (one proximity condition per block), resulting in a total of four experimental blocks over the two sessions.

Results and Discussion: Experiment 1

Figure 2 portrays the mean correlation (over subjects) of actual versus optimal response values for each of the 12 conditions. The figure illustrates that performance for the three display conditions (the three lines) was clearly ordered along the dimension of integrality (Integral rectangles tracked the optimal function with least error; bargraphs with greatest error). Statistical analysis revealed a significant main effect of integrality group, ($F(2,21)=4.76$, $p<.05$). In support of the principle of compatibility of display proximity, the data indicated that performance was significantly better for both of the more integral rectangle groups than for the more separable bargraph group, ($F(1,21)=4.32$, $p<.05$). No significant difference was found between the two rectangle conditions.

Secondly, a main effect of proximity condition, ($F(3,63)=6.80$, $p<.05$), revealed a significant benefit for cue presentations that were proximate in time (simultaneous vs. sequential presentation, $F(1,63)=9.37$, $p<.05$), as well as a significant decrement to performance under conditions of speed stress, ($F(1,63)=10.87$, $p<.05$). No significant benefit was found for manipulations of proximity in space (centered on screen vs. four distinct locations, $F(1,63)=2.38$, n.s.). It appears in Figure 2 as though manipulations of proximity had little influence on either of the rectangle groups and a large effect on the bargraph group (large benefits for proximity in space and large decrements under time stress). This interaction, however, was not statistically significant.

Finally it is important to note that memory for isolated, unintegrated attributes of reliability and diagnosticity was not harmed by the increasing integrality characteristic of either of the rectangular displays. Proportions of correct probes for the bargraph, rectangle and integral rectangle groups are .56, .51 and .65 respectively. In fact these differences suggest that there was actually a non-significant trend for the most integral condition to provide better memory of its isolated attributes.

There was one characteristic of the present paradigm that prevents us from concluding with certainty that there was no cost for attribute memory associated with the rectangle display. Since reliability and diagnosticity values for each cue were constant across trials and blocks, it is possible that some correct responses were a measure of long term memory for the values over a number of trials, rather than a measure of the subject's ability to recall information from short term memory about the immediately preceding trial. This possibility was tested in a follow-up experiment in which no contextual cues (with memorizable values) were present. This study also found no difference between displays in attribute memory, despite demonstrating the same pattern of significant differences in integration performance ($t(15)=2.0$, $p<.05$).

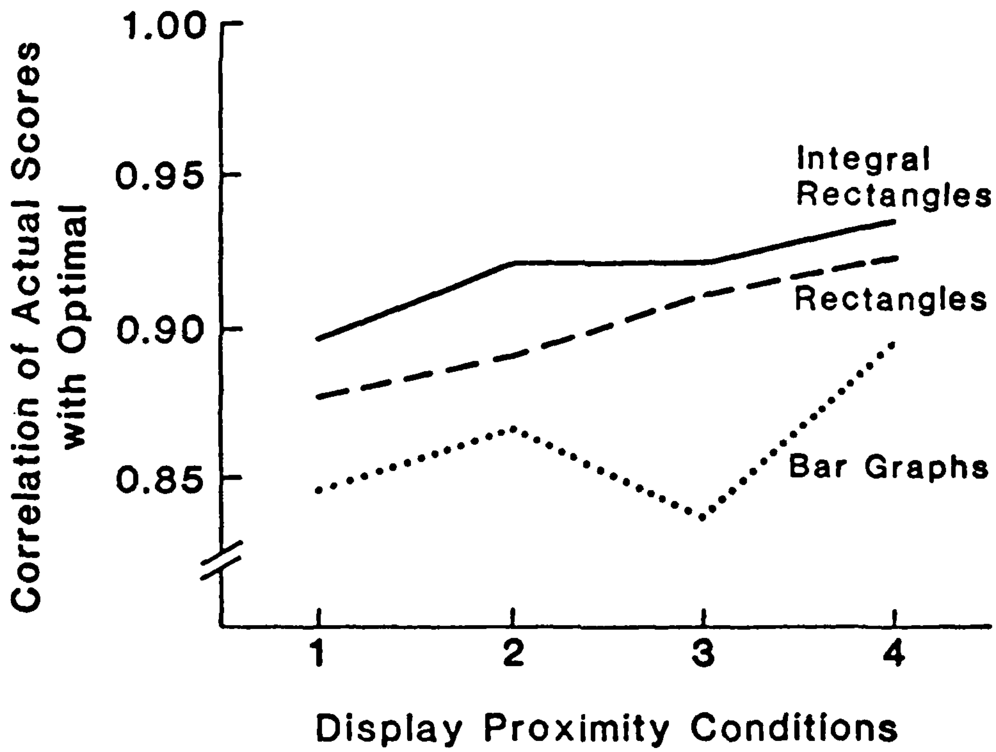

Figure 2. Mean correlations of actual with optimal scores plotted for each of the four proximity conditions for the three display integrality conditions (bargraphs, rectangles, & integral rectangles) (experiment 1).

Conclusion: Experiment 1

This study has contributed to the neglected human factors area of display formatting of probabilistic and decision-making information, an issue that increases in importance and prominence with the emerging technology in computer-based intelligent systems. While object-defined proximity compatability for integration tasks has been shown elsewhere (Carswell & Wickens, in press; Jones & Wickens, 1986), many of these benefits are also accompanied by a cost of processing individual attributes. Such costs have been demonstrated for the triangle display (Carswell & Wickens, in press; Goettl, Kramer & Wickens, 1986), and for the pentagon display (Casey & Wickens, 1986). An important new contribution of the present data is the finding of lack of cost to memory for isolated attributes with the rectangle display.

It appears that the emergent feature--the area--which provides relevant information for the integration task may facilitate encoding without disrupting memory for the raw features--height and width--that define the area. It may be that the rectangle display is a unique geometric object which satisfies these constraints, and provides an exception to the "no free lunch" principle, namely that you cannot obtain a benefit somewhere without incurring a cost elsewhere. While the present data do indeed support the conclusion of a rectangle integration benefit, without a cost to focussed attention on its attributes, such conclusion must be accepted with caution. Other research on rectangle perception for example by Fefoldy (1974), using a stimulus classification paradigm, has indeed found that height and width behave as "integral" dimensions; a cost is found in response time to one of its attributes as the other varies. It is possible that reaction time in the speeded classification task provided a more sensitive measure of focussed attention cost, than did attribute memory in the current paradigm. Finally, some indication of the potential usefulness of the rectangle display is the prevalence in the world of quantities that are directly defined by the product of two other meaningful quantities; for example, rate and time, pressure and temperature, or, as we have used here, reliability and diagnosticity.

EXPERIMENT 2: MULTICUE PROBABILITY LEARNING

Multicue probability learning (MCPL) is a close relation to the diagnostic judgment task examined in experiment 1. It is the task that confronts the analyst who must examine a set of predictor variables (e.g., test scores on an applicant) to assess their validity in predicting a criterion measure (e.g., success in the applicant's job). As such, it is an information integration task, and the normative model of performance is multiple regression. However, there are many field circumstances in which the operator may have neither the time nor the luxury to code data and carry out the regression analysis, but rather most form a rough estimate of the cue-criterion relation on the basis of a scan of visual data.

A study by Goldsmith and Schvanveldt (1981) provided a prototype for the current experiment. They displayed cue and criterion values for a set of observations either as bargraphs or as triangles, constructed by

connecting 3 radii at 120 degree angles, eminating from the center. Their experiment provided clear evidence for the advantage of the triangle display, across a set of different problem types. Consistent with the proximity-compatibility hypothesis, this advantage was attenuated in conditions that required some cues to be filtered.

The present experiment employed a task, similar to that used by Goldsmith and Schvanveldt, which required subjects to predict a criterion value given two or more displayed cue values. The subject's goal was to discover the formula that related the cues to the criterion. Since the task requires the subjects to combine the cues, an integral triangle display was predicted to improve the learning of the relation between the cues and the criterion. It was also predicted that when the task characteristics changed such that a subset of cues were to be selectively attended, the separable bargraph display would engender superior performance.

Method: Experiment 2

Twenty-four students with college level math were divided into four groups, each receiving a different combination of display and formula. Both displays presented three cue variables along the horizontal axis and the magnitudes of the cues along the vertical axis. The bargraph display represented the cues as the heights of three bars. In the integral triangle display the distance of each vertex of a triangle from the x-axis of the plot represented the magnitude of a cue variable; that is, the triangle was constructed by connecting the tops of each bargraph, and then deleting the bargraphs. Half the subjects within each display condition performed the task with an additive formula (criterion = (X + Y + Z) * 4), while the other half used a "noisy" formula (criterion = (X + Y + Z) * 4 + n, where n = any random number between -5 and +5).

Subjects were instructed that on each trial the three variables would be displayed for 10 seconds during which time they were to combine the values to predict a fourth, criterion, variable via an 11-key keypad. The correct criterion value as well as the subject's estimate was displayed as feedback to the subjects. The experiment was divided into two sessions, with each session consisting of three blocks of 120 trials. During the second session one cue was randomly selected to be irrelevant in the formula for each of the three blocks. Subjects were told that one value was to be ignored and were cued after each trial to "guess" which variable was irrelevant. An IBM-PC controlled the experiment and recorded response times along with the cue values, the criterion value, and the subjects' judgment on each trial. These data were used in a Brunswick lens analysis employing Tucker's (1964) equation:

$$r_a = G\ R_e\ R_s + C\sqrt{1-R_s'}\ \sqrt{1-R_e}$$

In this equation, r_a is the correlation between the subjects' judgments and the criterion values. It represents the subjects' achievement or performance. The variable G is the correlation between the linear predictions of the criterion values and the subjects' judgments from the cues. This component represents the subjects' knowledge of the task (i.e., mental model) independent of task uncertainty and knowledge execution. R_s

is the multiple correlation between the cue values and the subjects' judgments. This variable represents knowledge execution. R_e is the multiple correlation between the cue values and the actual criterion values. This factor represents task uncertainty. It is a measure of how accurately the cues predict the criterion and represents a limit on performance even if knowledge and knowledge execution are perfect. The final variable, C, is the correlation between the variance in the task and the subjects' judgments that is unaccounted for by G.

As can be seen from the equation, performance is a function of knowledge, execution of knowledge, and task uncertainty. Brunswik lens analysis allows for the separation of knowledge and performance. In the present study, both knowledge and performance were statistically analyzed. It should be noted that performance and knowledge are nearly equivalent for the additive formula, (X+Y+Z)*4. This is because R becomes 1 since there is no uncertainty or noise in the formula.

Results: Experiment 2

Knowledge scores (G) and performance scores (r_a) were transformed from correlations to z-scores using Fisher's z-transform equation:

$$Z = .5 \ln [(1 + r_a)/(1 - r_a)].$$

These data were then subjected to a 2(Session) x 3(Block) x 2(Formula) x 2(Display) mixed factors ANOVA. Session and Block factors served as repeated measures variables while Formula and Display factors represented between groups factors.

Analysis of the Z-scores provided partial support for the principle of proximity compatibility. The four way ANOVA performed on z-scores for G revealed an interaction between session and display ($F(1,20)=5.96$, $p<.05$). Figure 3 shows the z-scores relevant to this interaction. In this figure, z-scores are plotted along the y-axis while session is plotted along the x-axis. Bargraphs are represented with dashed lines while the triangle graphs are depicted as solid lines. The interaction was further broken down by session. The display effect in session 1 was not reliable. However, in session 2 the display effect was rather large ($F(1,20)=8.71$, $p<.01$). Thus, as Figure 3 suggests, there was no difference between the two display types when all three cues were relevant (session 1). When one cue became irrelevant and was to be selectively ignored (session 2), the bargraph display was superior to the triangle display. This demonstrates a limitation of integral displays and supports the second hypothesis of the proximity principle: When the task requires selective attention to a subset of the displayed variables, integral displays may interfere with performance (Carswell & Wickens, 1986).

There was also a strong interaction between session and formula ($F(1,20)=18.13$, $p<.01$). The additive formula showed improvement from session 1 to session 2 while knowledge about the noisy formula consistently dropped from session 1 to session 2. This interaction is undoubtedly due to the effect of the noise element (n) in the noisy formula. When only two

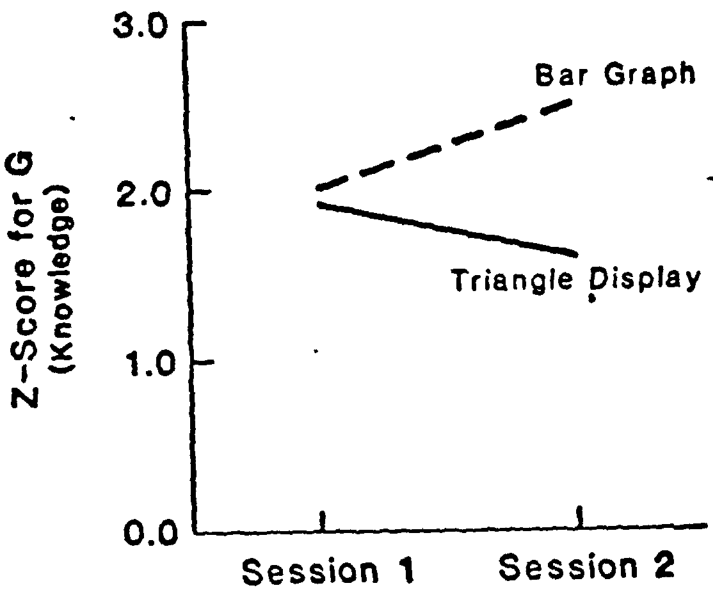

Figure 3. Mean G (knowledge) z-scores for the session by display interaction (experiment 2).

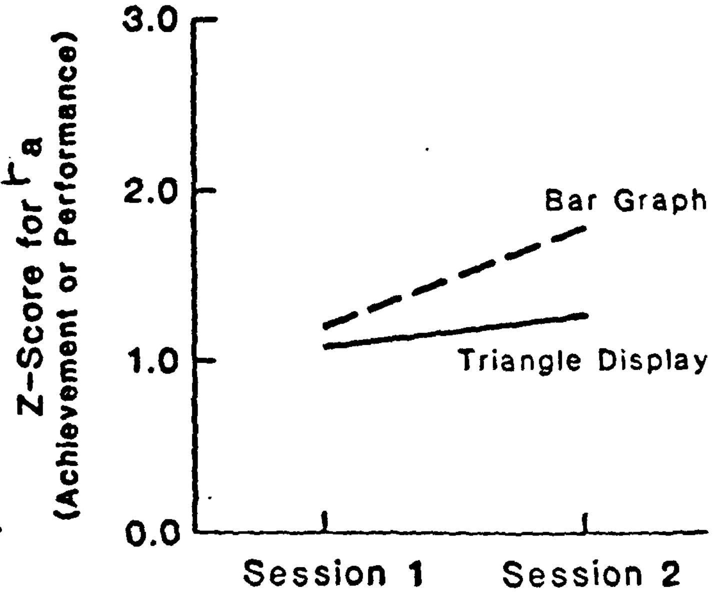

Figure 4. Mean r (performance) z-scores for the session by display interaction (experiment 2).

cues are relevant, the contribution of the random noise element is larger than when all cues are relevant and the task becomes more difficult.

When the data for the overall performance measure r_a were analyzed, the critical interaction between display and session did not achieve statistical reliability even though the data plotted in Figure 4, suggest a slight trend in the appropriate direction. Note the larger difference between bargraphs and triangle displays in session 2 as compared with session 1. The best explanation for why the performance data did not achieve reliability is that performance is based on several variables; knowledge, knowledge execution, task uncertainty, and random noise (factor C in Tucker's equation). Variance in any one of these factors could have accounted for the failure to find an interaction between display and session for performance measures. The Brunswick lens analysis allowed for knowledge to be examined independently of these other sources of variance. This suggests that the Brunswick lens analysis provides a better estimate of the relationship between display format and the mental representation of the task than measures of performance.

Finally, it should be noted that response time data were collected and subjected to a four-way ANOVA. This analysis revealed only practice effects; there were no effects related to display or formula. This is not surprising since the task was not performed under time stress: Subjects had 10 seconds to respond on each trial.

EXPERIMENT 3: MULTICUE PROBABILITY LEARNING AGAIN

Method: Experiment 3

This experiment utilized the multi-cue judgement task described in experiment 2. In addition to the bargraph display and the triangle display employed in the first experiment, a linegraph display was also used. This display plotted the cue values along the same axes used in the other two displays. Two lines connected the three cue values to form the linegraph. Three different formulas were used in this experiment. The first formula was the "noisy" formula used in experiment 2 ($C = (X+Y+Z)*4 + n$). The second formula, referred to as the correlational formula was quite similar to the first formula:

$$C = (X + XZ + Z)*4 + n, \text{ where } XZ = .75X + .25Z.$$

The middle cue, XZ, was correlated with X and Z rather than being a randomly selected value. This allowed us to examined the effects of correlation on object display benefits (or costs), a variable which other studies suggest may reduce the benefits (or increase the costs) of object displays. The third formula, referred to as the multiplication formula, was of the form:

$$C = (X * Y * Z) + 120)/9 + n.$$

This rather complex formula was chosen because it gives a similar range of values as the first two when all three cues are relevant. The three formulas were crossed with the three displays to form 9 groups of subjects.

Fifty-four subjects (27 males and 27 females) were recruited from the same population used in the first experiment. Six subjects were randomly assigned to each of the nine groups with the constraint that half the subjects in each group were males and half were females. The experimental procedure and method of data analysis were exactly the same as in experiment 2. In the second session, one of the three cues became irrelevant to the task and had to be ignored. Subjects were given no feedback as to which of the cues was irrelevant.

Results: Experiment 3

The results indicated that the effectiveness of the different displays was contingent upon the data structure (i.e. formula). A four-way ANOVA (2 sessions x 3 blocks x 3 displays x 3 formulas) was performed on z-score values of knowledge (factor G in Tucker's equation) and performance (factor r_a). These analyses revealed a marginal display x formula interaction ($F(4,45)=2.17$, $p=.0878$ for G; $F(4,45)=2.50$, $p=.0558$, for r_a). Figure 5 depicts the interaction revealed in knowledge data. Z-score values for knowledge are plotted along the y-axis and display type is represented along the x-axis. The plots for the three formulas are labeled as follows; A for the additive formula, C for the correlational formula, and M for the multiplicative formula. The simple main effects of display were evaluated for each formula separately. Display exerted a marginally reliable main effect for the additive formula ($F(2,15)=3.11$, $p=.0743$) but not for the correlational or the multiplicative formula. Figure 5 suggests that the additive formula was easier with the bargraph display than either the linegraph or the triangle displays. No display advantage was realized for the other formulas.

As in experiment 2, however, the processing demands were different in phase I, when all cues were to be integrated, and in phase II, when a selected subset were to be filtered out. Indeed the simple main effect of display type for the additive formula, indicated in Figure 5, was significant only in phase II when focused attention was required ($F(2,15)=2.97$; $p=0.082$). During complete integration in phase I, there was no difference over display type for any of the formulas. Therefore, consistent with experiment 2, when subjects were required to selectively attend to two of the three cues, the separable bargraph display was optimal.

Analysis of the performance rather than the knowledge data for experiment 3 revealed two interactions. The first was a session x formula interaction ($F2,45=23.04$; $p<.01$), with the additive display fairing relatively better in the focused attention sessions (II) and the multiplicative formula relatively better in the integration session (I). The correlation formula was consistently best in both sessions. The second was a marginal formulaXdisplay interaction ($F4,45=2.50$; $p=.09$), indicating that the multiplicative formula was affected by display type, while the additive and correlational formula were not. When this effect of display on the multiplicative formula was broken down by session, a strong display x session interaction was observed ($F2,15=7.20$; $p<.01$). The form of this interaction was somewhat predictable from the compatibility-proximity hypothesis. During session I (integration), the triangle display performed best, while during session II (filtering) this display performed most

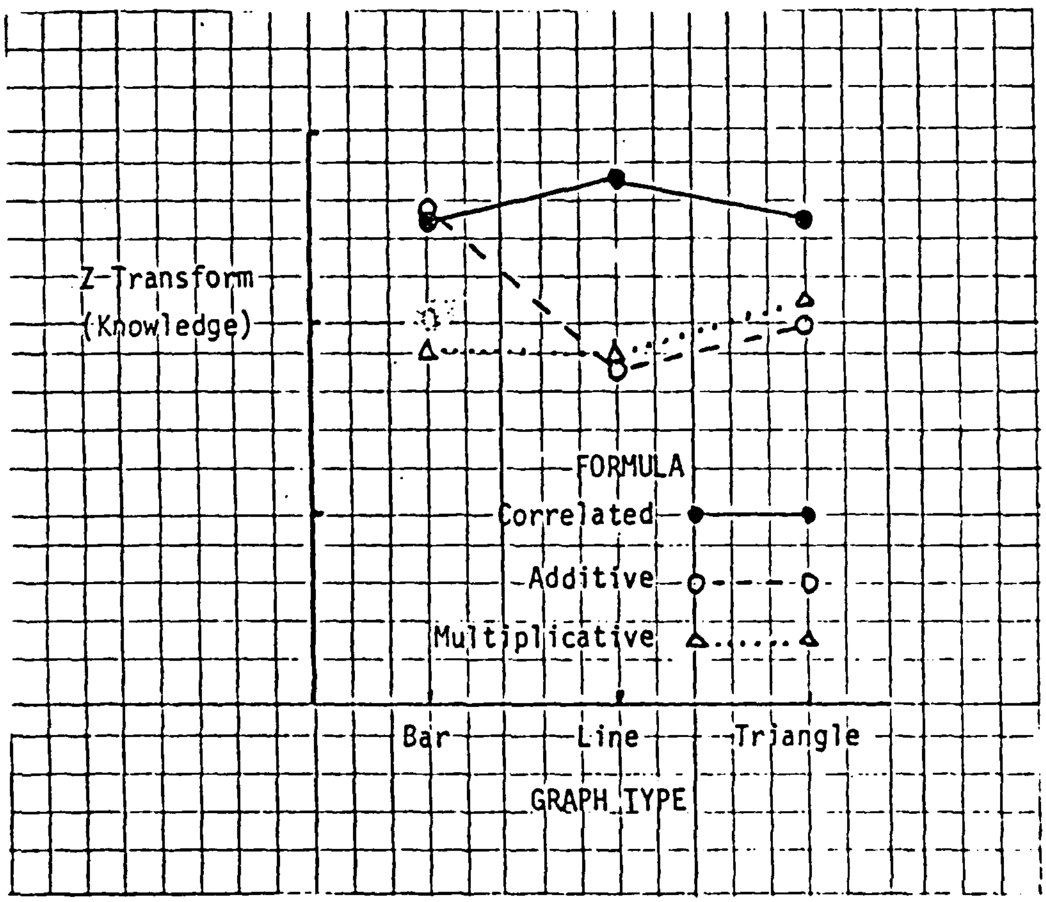

Figure 5. Effect of display type and formula on mean G (knowledge) z scores (experiment 3).

poorly. However, unexpectedly the worst display in session I, was not the most separate bargraph, but rather was the linegraph display, assumed a priori to impose an intermediate level of display proximity.

Discussion: Experiments 2 & 3

Taken together, experiments 2 & 3 are consistent with the pattern of data obtained by Goldsmith and Schvanveldt (1984), and demonstrate the applicability of the proximity principle in relatively simple tasks of multivariate data extrapolation. This principle consists of two hypotheses. First, when a task requires the integration of information the most optimal display presents the sources of information in close proximity (e.g., an integral display). Second, when the task does not require the integration of information or necessitates selective attention to a subset of the information, nonintegrated displays will be optimal. Experiment 2 supported the second hypothesis. In this experiment the interfering effects of an integral display were demonstrated in a multicue judgment task that required subjects to ignore one of the cues. Experiment 3 provided support for the first hypothesis when the multiplicative formula was examined. An advantage for the triangle display in the integration task (phase I), disappeared when filtering was required in phase II. Both experiments also demonstrated the importance of differentiating between knowledge and performance. Experiment 3 demonstrated the general argument that display effectiveness is contingent upon the underlying structure of the data. An interaction between display format and formula revealed that display format effected knowledge only when an additive relationship existed between independent cues and the criterion. When cues were correlated or combined multiplicatively, display format did not influence knowledge. However, display format affected performance only when the cues were combined multiplicatively.

One interesting facet of the data concerned the explicit manipulation of cue correlation in experiment 3. The broader pattern of data, presented in Figure 5 reveals a trend that is slightly at odds with other data collected and summarized in Wickens (1986). That is, the current data suggest that the presence of correlation will improve performance with an object display, relative to a more separated display. Such an effect, while consistent with early findings by Garner, concerning the affinity of integral dimensions (a relative of objectness) for redundant information, is at odds with more recent object display studies in our laboratory (Jones & Wickens, 1986) and others (Goldsmith & Schvanveldt, 1981). The source of discrepancy between these two trends remains to be clarified.

GENERAL DISCUSSION

The three experiments reported, reveal that object display benefits or costs, may be modified by the nature of the processing required of the displayed attributes in decision making tasks. This modification furthermore, is one of either an increased object benefit (experiment 1, performance data of experiment 3), or a decreased object cost (experiments 2 & knowledge data of 3), as the amount of integration is increased, or the amount of filtering is decreased. There remain however, a number of unresolved issues. The discrepancy for example between the presence of an object display advantage in Goldsmith and Schvanveldt's study, and its

general absence in experiments 2 & 3, cannot be readily explained. One possible source of the difference lies in the presence of noise in all of the criterion predictors in the current study, which was not present in Goldsmith and Schvanveldt's study. It is also worth noting that there was a fundamental difference in the way in which the two triangles were constructed in the two studies (extending radii from the center in Goldsmith & Schvanveldt; connected points on 3 vertical axes in the current studies). Thus, it is possible that differences in the physical construction of objects produce differences in emergent features of the displays that may serve to greater advantage in some paradigm than others (Pomerantz, 1981). For example, in Goldsmith and Schvanveldt's study, as in experiment 1 total display area was a salient variable that defined high values of all cues. In the current experiments 2 and 3 however, area was large only if the middle cue was disparate in value from the first and third cues.

Another unexplained finding was the difference in object-costs observed with independent and correlated cues in the knowledge of experiment 3. While puzzling, these effects of task type an the relative merits of object displays are sufficient to emphasize that display formats cannot be evaluated independently of the information they are intending to convey, nor the processing of this information that is intended. Further research is necessary to establish these principles relating the display with the cognitive domain. It is hoped that the proximity-compatibility principle will provide an initial step in this direction.

ACKNOWLEDGMENTS

We would like to thank Lisa Weinstein, Keith Glasch and Susan Wolsfeld for assistance in data collection for experiment 1; to thank Lisa Weinstein, Yili Liu, Rick Wermes, Othello Oandasan, and Mike Malay for assistance in running subjects for experiments 2 and 3; and Roger Marsh, Gunnar Seaburg and Todd Musleh for programming assistance.

REFERENCES

Anderson, N.H. (1981). Foundations of information integration theory. New York: Academic Press.

Boff, K.R., Kaufman, L. & Thomas, J.P. (Eds.). (1986). Handbook of Perception & Human Performance. New York: Wiley & Sons.

Carswell, M., & Wickens, C.D. (1986). Objections to objects. In L. Mark (Ed.), Trends in Ergonomics. (Vol. III).

Carswell, C. M., & Wickens, C. D. (in press). Information integration and the object display: An interaction of task demands and display superiority. Ergonomics.

Casey, E. J., & Wickens, C. D. (1986). Visual display representation of multidimensional systems. Cognitive Psychophysiology Laboratory Technical Report CPL 86-2/MDA903-83-K-0255. Department of Psychology, University of Illinois, Champaign, Illinois.

Cole, W.G. (1986). Medical cognitive graphics. Proceedings of the ACM-SIGCHI.

Fefoldy, G. (1974). Repetition effects in choice reaction time to multidimensional stimuli. Perception and Psychophysics, 15, 453-459.

Goettl, B.P., Kramer, A.F., & Wickens C.D. (1986). Display format and the perception of numerical data. Proceedings of the Human Factors Society 30th Annual Mtg.

Goldsmith, T.E., & Schvanveldt, R.W. (1981). Facilitating multiple-cue judgments with integral information displays. In J.C. Thomas and M.L. Schneider (Eds.), Human factors in computing systems. Norwood, NJ: Ablex Publishing Co.

Johnson, E.M., Cavanagh, R.C., Spooner, R.L., & Samet, M.G. (1973). Utilization of reliability measurements in Bayesian inference: Models and human performance. IEEE Transactions on Reliability, 22, 176-183.

Jones, P.M., & Wickens, C.D. (1986). The display of multivariate information: The effects of auto and cross-correlation, reliability, and heterogeneity. ARI CPL Technical Report No. CPL-86-5.

Lopes, L.L. (1981). Averaging rules & adjustment processes: The role of averaging in inference. (Technical Report WHIPP 13). Madison, WI: Wisconsin Human Information Processing Program.

Pitz, G.F. & Sachs, N.J. (1984). Judgment and decision: Theory and application. Annual Review of Psychology, 35, 139-163.

Pomerantz, J.R. (1981). Perceptual organization in information processing, In R. Kubovy & J.R. Pomerantz (Eds.), Perceptual organization, Hillsdale, N.J.: Erlbaum.

Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd. Psychological Review, 71, 528-530.

Wickens, C.D. (1986). The object display: Principles and a review of experimental findings (CPL Tech. Rep. CPL-86-6). Champaign, IL: University of Illinois, Department of Psychology.

Wickens, C.D. & Scott, B.D. (1983). A comparison of verbal and graphical information presentation in a complex information integration task. (Technical Report EPL-83-1/ONR-83-1), Engineering-Psychology Research Laboratory, University of Illinois at Urbana-Champaign.